APPLICATION

FOR

UNITED STATES LETTERS PATENT

APPLICANT NAME: Martin et al.

TITLE: LOWER OVERHEAD SHARED CACHE INVALIDATIONS

DOCKET NO.: RSW920030299US1

INTERNATIONAL BUSINESS MACHINES CORPORATION

CERTIFICATE OF MAILING UNDER 37 CFR 1.10

I hereby certify that, on the date shown below, this correspondence is being deposited with the United States Postal Service in an envelope addressed to Mail Stop Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 as "Express Mail Post Office to Addressee" Mailing Label No. EV393299659US

on March 15, 2004

Beverly Kehoe Shea
Name of person mailing paper

Signature

03/15/2004
Date

LOWER OVERHEAD SHARED CACHE INVALIDATIONS

Background of the Invention

1. Field of the Invention

[0001] The present invention generally relates to caching. More particularly, the present invention provides a method, system and program product for reducing the overhead of cache invalidations in a shared cache (e.g., a cache cluster) by transmitting a hashed code of a key to be invalidated rather than the entire key.

2. Related Art

[0002] Caching is used in many systems to reduce access time to data. In a clustered environment, caches are often distributed among multiple systems in a cluster. A group of caches in a cluster that all cache the same data are called "peers." To provide cache consistency among peers, caches may send invalidation messages to their peers when an object associated with a key is modified or deleted from the cache.

[0003] Sending an invalidation to peers in current shared cache systems involves sending the entire key of the object to be invalidated. In Java, as shown in the flow diagram of FIG. 1, this is typically done by serializing the key (step S1), transmitting the serialized version of the key to each peer (step S2), and deserializing the received serialized version of the key at each peer (step S3). The key received by each peer is subsequently discarded from the cache (step S4). Unfortunately, there is a processing

overhead in the serialization/deserialization of the key. Further, the entire key must be transmitted, which consumes extra network bandwidth. As such, there is a need to reduce the processing and bandwidth overhead of cache invalidations in a shared cache.

Summary of the Invention

[0004] In general, the present invention provides a method, system, and program product for reducing the overhead of cache invalidations in a shared cache by transmitting a hashed code of a key to be invalidated rather than the entire key.

[0005] A first aspect of the present invention is directed to a method for shared cache invalidation, comprising: hashing a key corresponding to an object in a first cache that has been modified or deleted to provide a hashed code of the key, wherein the first cache forms part of a shared cache; transmitting the hashed code of the key to other caches in the shared cache; comparing the hashed code of the key with entries in the other caches; and dropping any keys in the other caches having a hash code the same as the hashed code of the key.

[0006] A second aspect of the present invention is directed to a method for shared cache invalidation, comprising: hashing a key corresponding to an object in a cache that has been modified or deleted to provide a hashed code of the key, wherein the cache forms part of a shared cache; and transmitting the hashed code of the key to other caches in the shared cache for invalidation processing.

[0007] A third aspect of the present invention is directed to a method for shared cache invalidation, comprising: receiving at a cache a hashed code of a key, wherein the hashed code of the key corresponds to an object that has been modified or deleted in another cache; and performing invalidation processing on the cache based on the received hashed code of the key.

[0008] A fourth aspect of the present invention is directed to a system for shared cache invalidation, comprising: a system for hashing a key corresponding to an object in a cache that has been modified or deleted to provide a hashed code of the key, wherein the cache forms part of a shared cache; and a system for transmitting the hashed code of the key to other caches in the shared cache for invalidation processing.

[0009] A fifth aspect of the present invention is directed to a system for shared cache invalidation, comprising: a system for receiving at a cache a hashed code of a key, wherein the hashed code of the key corresponds to an object that has been modified or deleted in another cache; and a system for performing invalidation processing on the cache based on the received hashed code of the key.

[0010] A sixth aspect of the present invention is directed to a program product for shared cache invalidation stored on a recordable medium, which when executed, comprises: program code for hashing a key corresponding to an object in a cache that has been modified or deleted to provide a hashed code of the key, wherein the cache forms part of a shared cache; and program code for transmitting the hashed code of the key to other caches in the shared cache for invalidation processing.

[0011] A seventh aspect of the present invention is directed to a program product for shared cache invalidation stored on a recordable medium, which when executed, comprises: program code for receiving at a cache a hashed code of a key, wherein the hashed code of the key corresponds to an object that has been modified or deleted in another cache; and program code for performing invalidation processing on the cache based on the received hashed code of the key.

Brief Description of the Drawings

[0012] These and other features of this invention will be more readily understood from the following detailed description of the various aspects of the invention taken in conjunction with the accompanying drawings in which:

[0013] FIG. 1 depicts a flow diagram of a cache invalidation method in accordance with the prior art, in which an entire key is serialized/deserialized.

[0014] FIG. 2 depicts a flow diagram of a cache invalidation method using a hashed code of a key, in accordance with the present invention.

[0015] FIG. 3 depicts an illustrative shared cache system comprising a plurality of application servers, in which the method of the present invention may be utilized.

[0016] FIG. 4 illustrates an example of a computerized system capable of implementing the present invention.

[0017] The drawings are merely schematic representations, not intended to portray specific parameters of the invention. The drawings are intended to depict only

typical embodiments of the invention, and therefore should not be considered as limiting the scope of the invention. In the drawings, like numbering represents like elements.

Detailed Description of the Invention

[0018] As indicated above, the present invention provides a method, system, and program product for reducing the overhead of cache invalidations in a shared cache by transmitting a hashed code of a key to be invalidated. It is assumed for the purposes of this description that the reader has an understanding of caches, cache indexing, shared cache systems, cache invalidation techniques, Java, and hashing commensurate with one skilled in the art. Accordingly, a detailed description of these topics is not provided herein.

[0019] Referring now to FIG. 2, there is illustrated a flow diagram of a cache invalidation method using a hashed code of a key in accordance with the present invention. In step S10, a key corresponding to a cached object that has been modified or deleted in a cache is hashed using a hashing algorithm. Any suitable hashing algorithm now known or later developed may be utilized in the practice of the present invention. In step S1 of FIG. 1, the serialized key may have a size of tens of bytes to hundreds of bytes. The hashed code of the key of the present invention produced in step S10 of FIG. 2, however, may typically have a length of four or eight bytes. As such, the hashed code of the key of the present invention has a much shorter length than a corresponding serialized key, and a corresponding lower bandwidth consumption when transmitted.

[0020] In step S11 of FIG. 2, the hashed code of the key is transmitted to each peer of the cache. The manner of transmission (e.g., via a hard-wired connection, via the Internet, etc.) is dependent upon the specific configuration of the shared cache.

[0021] The present invention takes advantage of the fact that caches commonly index entries using hash codes. In particular, in step S12 of FIG. 2, upon receiving the hashed code of the key, each peer cache drops all keys having the same hash code. Since two equivalent keys must have the same hash code, each peer cache is guaranteed of invalidating the desired key. Thus, by using the hashed code of the key, the process of locating matching keys in each peer cache is greatly simplified. In addition, processing requirements are reduced relative to the serializing/deserializing process of the prior art.

[0022] There is a small possibility that extra keys may also be invalidated. However, this is extremely unlikely given the typical range of hash values compared to typical cache sizes. That is, cache keys can typically be hashed to a numeric hash code having a range much larger than the typical cache size. For example, on 32-bit hardware, Java objects can generate a hash code with a range of approximately four billion options (i.e., 4×10^9), while caches often contain only thousands or tens of thousands of objects. Regardless, even if an extra key is accidentally invalidated, the cache still operates correctly: caches can arbitrarily and capriciously discard entries whenever they wish.

[0023] An illustrative shared cache system 100 in which the method of the present invention may be utilized is illustrated in FIG. 3. In this example, the shared cache system 100 comprises three application servers 102a, 102b, and 102c. Although

shown as including three application servers 102a, 102b, and 102c, it will be apparent that the shared cache system 100 may comprise any plurality of application servers. Also, it should be understood that the architecture shown in FIG. 3 is illustrative only and will likely include other known components not shown. Further, it will be appreciated that the present invention can be used to lower the processing and bandwidth overhead of shared cache invalidations in many other types of systems that include a shared/distributed cache including, for example, web servers, databases, etc.

[0024] Each application server 102a, 102b, and 102c comprises a cache 104a, 104b, and 104c, and a cache controller 106a, 106b, and 106c, respectively. The application servers 102a, 102b, and 102c can be interconnected (as shown in phantom) in any known or later developed manner. In this example, it is assumed that an object stored in the cache 104a of application server 102a has been modified or deleted. In response thereto, the cache controller 106a accesses and hashes the key of the modified/deleted object to provide a hashed code of the key. The cache controller 106a then transmits the hashed code of the key to the other cache peers (e.g., application servers 102b and 102c) in the shared cache system 100. Upon receipt of the hashed code of the key from the application server 102a, the cache controller 106b of application server 102b compares the hashed code of the key with the entries in the cache 104b and drops all keys having the same hash code. Likewise, upon receipt of the hashed code of the key from the application server 102a, the cache controller 106c of application server 102c compares the hashed code of the key with the entries in the cache 104c and drops all keys having the same hash code.

[0025] Referring now to FIG. 4, a computer system 200 capable of implementing the method of the present invention is illustrated in detail. The computer system 200 represents, for example, a portion of a shared cache system, such as one of the application servers 102 shown in FIG. 3. As shown, the computer system 200 generally comprises a central processing unit (CPU) 202, memory 204, bus 206, input/output (I/O) interfaces 208, external devices/resources 210, storage unit 212, and a shared cache 214. CPU 202 may comprise a single processing unit, or may be distributed across one or more processing units in one or more locations. Memory 204 may comprise any known type of data storage and/or transmission media, including magnetic media, optical media, random access memory (RAM), read-only memory (ROM), a data cache, a data object, etc. Moreover, similar to CPU 202, memory 204 may reside at a single physical location, comprising one or more types of data storage, or may be distributed across a plurality of physical systems in various forms.

[0026] I/O interfaces 208 may comprise any system for exchanging information to/from an external source. External devices/resources 210 may comprise any known type of external device, including speakers, a CRT, LCD screen, handheld device, keyboard, mouse, voice recognition system, speech output system, printer, monitor/display, facsimile, pager, etc. Bus 206 provides a communication link between each of the components in computer system 200 and likewise may comprise any known type of transmission link, including electrical, optical, wireless, etc.

[0027] Storage unit 212 can be any system capable of providing storage for information necessary for the practice of the present invention. As such, storage unit 212

may reside at a single physical location, comprising one or more types of data storage, or may be distributed across a plurality of physical systems in various forms. In another embodiment, storage unit 212 may be distributed across, for example, a local area network (LAN), wide area network (WAN) or a storage area network (SAN) (not shown).

[0028] Shown in memory 204 of computer system 200 is a cache controller 220, which may be provided as computer program product. The cache controller 220 is provided to determine when an object in the shared cache 214 has been modified or deleted. In response thereto, the cache controller 220 accesses the key of the modified or deleted object and hashes the key to provide a hashed code of the key. The cache controller 220 then transmits the hashed code of the key to the other cache peers in the shared cache system. Upon receipt of a hashed code of the key from any of the other cache peers in the shared cache system, the cache controller 220 compares the received hashed code of the key with the entries in the shared cache 214 and drops all keys having the same hash code.

[0029] If a shared cache is going to use the present invention in an environment where the same key may give different hash codes on different nodes in the cluster, then the invention may not be applicable. One example where this can occur is in a cluster of Java applications where different nodes run on different hardware platforms; in that case, the hash code for a given key may be different. However, a cache can be configured to include a detection system (e.g., hash code detection system 222, FIG. 4) to automatically detect this for each type of key inserted into the cache by computing the hash code for the

key, and then sending the key and its hash code to all its peer caches, which also hash the same key. If the peer caches also compute the same hash code, then the invention can safely be applied for this type of key. If different hash codes result, then this invention is not applicable for the specified key type. This could be done once per peer in the cluster, to detect if this invention is safe to use.

[0030] It should be understood that the present invention can be realized in hardware, software, or a combination of hardware and software. Any kind of computer/server system(s) - or other apparatus adapted for carrying out the methods described herein - is suited. A typical combination of hardware and software could be a general purpose computer system with a computer program that, when loaded and executed, carries out the respective methods described herein. Alternatively, a specific use computer, containing specialized hardware for carrying out one or more of the functional tasks of the invention, could be utilized. The present invention can also be embedded in a computer program product, which comprises all the respective features enabling the implementation of the methods described herein, and which - when loaded in a computer system - is able to carry out these methods. Computer program, software program, program, or software, in the present context mean any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function either directly or after either or both of the following: (a) conversion to another language, code or notation; and/or (b) reproduction in a different material form.

[0031] The foregoing description of the preferred embodiments of this invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and obviously, many modifications and variations are possible. Such modifications and variations that may be apparent to a person skilled in the art are intended to be included within the scope of this invention as defined by the accompanying claims.